Estimating Stock Price Predictability Using Genetic Programming

Minglei Duan
Dept. of Electrical and Computer Engineering
Marquette University
PO Box 1881, Milwaukee, WI 53201-1881

Richard J. Povinelli
Dept. of Electrical and Computer Engineering
Marquette University
PO Box 1881, Milwaukee, WI 53201-1881

A new method that quantifies the predictability of a stock's price is presented. This new method overcomes some problems with previous approaches to stock price predictability. The new method shows which stocks are more predictable, and hence can help in making investment decisions. Genetic Programming is used as the modeling tool. Preliminary experiments are conducted to show the advantages of this method.

Time series predictability is a measure of how well future values of a time series can be forecasted. Measuring the predictability of a time series is important because it can tell whether a time series can be predicted before making a prediction. Therefore prediction of time series with low predictability, such as a random walk time series, can be avoided. A good metric for time series predictability also provides a measure of confidence in the accuracy of a prediction. This is especially helpful to minimize the risk when making an investment decision.

An h-metric was introduced by Kaboudan [1], which measures the probability that a series is GP-predictable using Genetic Programming [2]. By design, the computed metric should approach zero for a complex signal that is badly distorted by noise. Alternatively, the computed metric should approach one for a time series with low complexity and strongly deterministic signal.

This metric is based on comparing two outcomes: the best fit model generated from a single data set before shuffling with the best fit model from the same set after shuffling. The shuffling process is done by randomly scrambling the sequence of an observed data set using Efron's bootstrap method. Specifically, the unexplained variations, which are measured by the sum of squared error before and after shuffling of a time series are compared.

While applying the h-metric to estimate stock price predictability, two main problems have been observed. First, the value of the metric largely depends on the length of the time series. The source of this effect is mainly due to the nonstationarity of stock price time series, and the nonstationarity becomes more evident as the sample size increases. The second problem is a derivation of the first one. For a long-term stock series, the value of the h-metric will be distributed in a very narrow range. Hence, it largely reduces the resolution of the metric.

These problems are resolved by using a shifting window, so that the h-metric can be estimated on the same sample size, as long as the window length is fixed. By changing the window length, the range of the metric can be adjusted. The overall predictability of a time series can be estimated by calculating the average h over all windows.

In order to test the new metric, it is applied to three different kinds of time series: a deterministic time series (Mackey-Glass map), a stock price time series, and a random walk time series. It can be seen clearly from the experiments that different kinds of times series yield significantly different results (Table 1).

Table 1: Experimental results

Time series	**Predictability**
Mackey-Glass	0.996
Random Walk	0.140
GE Stock	0.485

This method has been shown to be able to distinguish stock price time series and random walk time series effectively. Future work needs to be done comparing more different stocks to test whether the metric can quantify the predictability accurately. This method may be helpful in making investment decisions.

References

[1] Kaboudan, M. 1998. A GP approach to distinguish chaotic from noisy signals. *Genetic Programming 1998: Proceedings of the Third Annual Conference*, San Francisco. CA: Morgan Kaufmann, pp. 187-192

[2] Koza, John 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA: The MIT Press.

[3] Kaboudan, M. Genetic Programming Prediction of Stock Prices, *Computational Economics*, to appear.